Double-Talk and Path Change Detection Using A Matrix of Correlation Coefficients

Background Of The Invention

1.   Field of Invention

[0001] This invention relates to a method of detecting double-talk and path changes in echo cancellation systems. Echo cancellation is used extensively in telecommunications applications to recondition a wide variety of signals, such as speech, data transmission, and video.

2.   Description of Related Art

[0002] The search for an effective echo cancellation procedure has produced several different approaches with varying degrees of complexity, cost, and performance. A traditional approach to echo cancellation uses an adaptive filter of length L, where L equals the number of samples necessary to extend to just beyond the duration of the echo. Typically, the adaptive filters contain either 512 or 1024 taps. At the standard telephone bit rate of 8000 samples per second, this provides the ability to adapt to echo paths as long as 64 ms and 128 ms, respectively.

[0003] The computational requirements of an adaptive filter are proportional to L for the popular LMS (Least Mean Squares) class of algorithms, and proportional to L^2 or higher for algorithms such as RLS (Recursive Least Squares). More robust algorithms (like RLS) have greatly improved convergence characteristics over LMS methods, but the L^2 computational load makes them impractical with current technology. For this reason, the LMS algorithm (and its variants) tends to remain the algorithm of choice for echo cancellation.

[0004] Practical echo cancellation devices must provide some means of avoiding divergence from double-talk. The double-talk condition arises when there is simultaneous transmission of signals from both sides of the echo canceller due to the presence of near-end speech in addition to the echo. Under such circumstances, the return echo path signal, S_{IN} (see Figure 1), contains both return echo from the echo source signal, and a double-talk signal. The presence of a double-talk signal will prevent an LMS-based echo canceller from converging on the correct echo path. It will also cause a pre-converged

1

echo canceller to diverge to unpredictable states. Following divergence, the echo canceller will no longer cancel the echo, and must reconverge to the correct solution. Such behaviour is highly unacceptable, and is to be avoided in actual devices. Some means of detecting double-talk must therefore be implemented. To prevent divergence, the LMS filter coefficients are typically frozen during the presence of double-talk.

[0005] Detecting double-talk quickly and reliably is a notoriously difficult problem. Even a small amount of divergence in a fully converged LMS filter will result in a significant increase in the residual echo level. The use of a fast and reliable double-talk detector is crucial to maintain adequate subjective performance.

[0006] The simplest, and perhaps most common, method for detecting double-talk is to use signal levels. The echo path typically contains a minimum amount of loss, or reduction, in the return signal. This quantity is often referred to as the Echo Return Loss, or ERL. In most systems, this is assumed to be at least 6 dB. In other words, the return signal S_{IN} will be at a level which is at least 6 dB lower than R_{OUT} provided that there is no double-talk. In the presence of double-talk, the level at S_{IN} often increases so that it is no longer 6 dB lower than R_{OUT}. This condition provides a simple and convenient test for double-talk.

[0007] The problem with this approach is that the double-talk detector must have an accurate estimate of the echo path ERL in order to determine if the level at S_{IN} is too high. However, precise knowledge of the ERL is generally not available. If the ERL estimate is too high, the double-talk detector may trigger unnecessarily. Conversely, it may not trigger at all if the ERL estimate is too low.

[0008] Another problem with this technique is that it will only reliably detect high-level double-talk. If the double-talk signal is at a much lower level than the echo source signal, low-level double-talk occurs. Under this condition, the increase in the level of S_{IN} is usually very small. The double-talk detector may fail to trigger, but noticeable divergence in the LMS filter can still occur.

[0009] To detect low-level double-talk, the level of the residual echo signal (S_{OUT}) is often monitored. If no double-talk or background noise is present, and the LMS filter is fully converged, S_{OUT} can be as much as 40 dB lower than R_{OUT}. Assuming that the echo

2

path remains constant, any increase in S_{OUT} will likely be due to double-talk. Of course, if the echo path does change, it will be mistaken for double-talk. So if this method is used, a separate path change detection algorithm must be employed. A unified approach would be simpler and preferred.

5   [0010] Correlation is a statistical function which is commonly used in signal processing. It can provide a measure of the similarity between two signals (cross-correlation), or a single signal and time-shifted versions of itself (autocorrelation). The use of correlation for double-talk detection per se is known. Several patents exist for correlation-based double-talk detection, including US5646990, US5526347 and US5193112. The

10   correlation-based approaches taken in prior-art methods generally involve the calculation of a single cross-correlation coefficient, usually between R_{IN} and S_{IN}. The problem with this technique is that the degree of correlation can vary widely with different signals and echo paths. This makes it very difficult to set thresholds on the correlation coefficient in order to determine what state the echo canceller is in.

15   **Summary Of The Invention**

[0011] A process has been developed which generates matrix coefficients using zero-lag auto and cross-correlations from signals commonly found in echo cancellers. Double-talk and path changes are then detected using matrix operations such as determinants, eigendecompositions, or singular value decompositions (SVDs).

20   [0012] The correlations between various signals in an echo canceller will change depending on what state the echo canceller is in, i.e. if it is converged, unconverged, or in double-talk. By arranging the various correlations in appropriate matrix form, key information about the state of the echo canceller can be extracted by performing various matrix operations. The preferred operation is to take the determinant , but

25   eigendecompositions and singular value decompositions (SVDs) can also be used. A novel aspect of the invention is the formulation of a matrix using various correlation coefficients, and the subsequent analysis of this matrix to determine the state of the echo canceller.

[0013] Accordingly the present invention provides a method of detecting double-talk and

30   path changes in an echo cancellation system, comprising generating a correlation-based

matrix of signals in said echo cancellation system; and analyzing said correlation-based matrix to identify double-talk and path changes occurring in said system.

[0014] In the preferred embodiment, the correlation-based matrix is generated using the return echo signal (S_{IN}) and the output of an LMS adaptive filter.

[0015] The invention provides a correlation-based matrix is generated using zero-lag auto and cross-correlations of signals commonly found in echo cancellers.

[0016] Double-talk and path changes are detected by analysis of the correlation-based matrix. Possible analysis techniques include condition numbers, determinants, eigendecompositions, and singular value decompositions.

[0017] In the preferred embodiment, determinants are used to detect double-talk and path changes.

[0018] The invention can be implemented using either the time-domain or frequency-domain in a digital signal processor using conventional digital signal processing techniques.

[0019] The invention also provides a double-talk and path change detector, comprising a processing element generating a correlation-based matrix of signals in said echo cancellation system; and a processing element for analyzing said correlation-based matrix to identify double-talk and path changes occurring in said system.

Brief Description Of The Drawings

[0020] The invention will now be described in more detail, by way of example only, with reference to the accompanying drawings, in which;-

Figure 1 is a schematic diagram of an echo canceller using LMS Adaptive Filtering; and

Figure 2a is a plot showing the value of det [R] under normal convergence;

Figure 2b is a plot showing the value of det [R] with a path change; and

Figure 2c is a plot showing the value of det [R] with double-talk.

4

Detailed Description Of The Invention

[0021] The layout of a typical LMS-based echo canceller is shown in Figure 1. It contains two signals, travelling along a "send" path and a "receive" path. The echo source signal enters the echo canceller as R_{IN} and leaves as R_{OUT}. The send path input, S_{IN}, consists of a double-talk signal (if present) plus the echo source signal after it has travelled along the echo path. By estimating the echo path, a synthetic echo signal can be generated to cancel the echo in the send path. The echo cancelled signal leaves as S_{OUT}.

[0022] The LMS filter attempts to cancel the echo by adjusting itself to suppress the output signal at S_{OUT}. Obviously, if S_{IN} contains components other than echoed speech from the echo source, the LMS filter will not converge to the correct solution; hence the need for double-talk detection.

[0023] The preferred embodiment of the algorithm for this patent uses the Normalized-LMS (N-LMS) algorithm. Mathematically, the adaptive filter tap-weight update procedure for the N-LMS algorithm consists of the following three equations

$$\hat{d}[n] = \hat{w}^H[n]u[n]$$

$$e[n] = d[n] - \hat{d}[n]$$

$$\hat{w}[n+1] = \hat{w}[n] + \frac{\mu}{a + \|u[n]\|^2}u[n]e[n]$$

[0024] where

$u[n] = R_{IN}$ = echo source signal

$\hat{w}[n]$ = LMS filter coefficients

$d[n] = S_{IN}$ = desired LMS output (echo + double-talk)

$\hat{d}[n]$ = LMS output (estimated echo)

$e[n] = S_{OUT}$ = LMS error signal

μ = LMS step-size parameter

a = A small constant (provides numerical stability).

5

[0025] The location of these signals in the echo canceller is also shown in Figure 1. The N-LMS algorithm well known to persons skilled in the art and a more detailed treatment is readily available in most adaptive filtering texts. See, for example, S. Haykin, Adaptive Filter Theory, Prentice-Hall, Upper Saddle River, NJ (1996), the contents of which are herein incorporated by reference.

[0026] One of the key parameters in the N-LMS algorithm is the LMS step-size parameter μ. This parameter controls both the speed and accuracy of convergence. The larger μ is, the faster the algorithm will converge on the echo path, but the less accurate the steady-state solution will be. To guarantee convergence of the N-LMS algorithm, μ must be less than 2.

[0027] A common technique is to adjust the value of μ based on the state of the echo canceller. In an unconverged state (such as at start-up, or following a path change), it is desirable to use a large value for μ to permit rapid initial convergence. Once the LMS filter has achieved a reasonable degree of convergence, μ can be reduced. This not only allows for a slightly more accurate solution (and therefore more cancellation), but it will also slow potential divergence from double-talk. To stop adaptation altogether, μ can simply be set to zero. The double-talk and path change detectors can therefore control the operation of the LMS filter by varying the value of μ.

[0028] A double-talk detection algorithm in accordance with a preferred embodiment of the invention that is designed to work in conjunction with the echo canceller described illustrated in Figure 1 will be described. This is implemented in a digital signal processor.

[0029] Consider two signals, $X_0[n]$ and $X_1[n]$ generated by a linear combination of two real-valued source signals, $S_0[n]$ and $S_1[n]$. Mathematically, this mixing process may be described as

$$X_i = H_{i,0} \cdot S_0 + H_{i,1} \cdot S_1 ,$$

[0030] where $H_{i,j}$ are the mixing coefficients. In matrix form, this may be written as

$$X = H \cdot S$$

[0031] where

6

$$\mu \quad X = \begin{bmatrix} X_0 \\ X_1 \end{bmatrix}, H = \begin{bmatrix} H_{0,0} & H_{0,1} \\ H_{1,0} & H_{1,1} \end{bmatrix} \text{ and } S = \begin{bmatrix} S_0 \\ S_1 \end{bmatrix}$$

[0032] A matrix R is defined as

$$\mu \quad R = E\begin{bmatrix} XX^T \end{bmatrix}$$

[0033] where $E[\ldots]$ is the statistical expectation operator. R may be expanded in two ways

$$\mu \quad R = E\begin{bmatrix} X_0 X_0^T & X_0 X_1^T \\ X_1 X_0^T & X_1 X_1^T \end{bmatrix}$$

$$\mu \quad = E\begin{bmatrix} HSS^T H^T \end{bmatrix}$$

[0034] From the first expansion, it is apparent that the diagonal terms in R are the zero-lag autocorrelations of $X_0[n]$ and $X_1[n]$ and that both off-diagonal terms correspond to the zero-lag cross-correlation between $X_0[n]$ and $X_1[n]$. Hence, R is a symmetric, correlation-based matrix.

[0035] From the second expansion, we see that if H is full-rank, then R will also be full-rank if $S_0[n]$ and $S_1[n]$ are both non-zero and uncorrelated. In most cases, a sufficient condition for this is that $S_0[n]$ and $S_1[n]$ are different signals from different sources.

[0036] The way in which the matrix can be used to perform double-talk and path change detection will now be explained. First, suppose we generate the signal mixtures in using convolutions:

$$\mu \quad X = H \otimes S$$

[0037] Now the terms in the mixing matrix can be vectors. We further impose the condition that H have the following form:

$$\mu \quad H = \begin{bmatrix} H_{0,0} & 1 \\ H_{1,0} & 0 \end{bmatrix}$$

7

[0038] With $\mathbb{H}$ defined in this way, it is now possible to connect the terms in the preceding equations with the parameters available in the echo canceller layout shown in Figure 1. Let

μ S_0 = echo source signal = R_{IN} = $u[n]$

μ S_1 = double-talk signal

μ $H_{0,0}$ = echo path

μ $H_{1,0}$ = LMS filter coefficients = $\hat{w}[n]$

[0039] With these definitions, it is apparent that

μ $X_0 = H_{0,0} \otimes S_0 + S_1 = S_{IN} = d[n]$

μ $X_1 = H_{1,0} \otimes S_0 = \hat{d}[n]$

[0040] The question of what happens to $\mathbb{R}$ under the various states of echo canceller operation will now be examined.

[0041] Case 1: Unconverged, no double-talk

[0042] If the LMS filter is in an unconverged state, $H_{0,0} \neq H_{1,0}$. This situation occurs when the echo canceller is first started, or following a major echo path change. Since the LMS filter does not contain an accurate echo path estimate, $X_0 \neq X_1$, and $\mathbb{R}$ will be full rank (unless $H_{1,0} = 0$, but this condition is usually temporary) with a very low condition number. See, for example, G. H. Golub and C. F. Van Loan, Matrix Computations, 3rd ed., Johns Hopkins University Press, Baltimore, MD (1996). ($\kappa \sim 10^1$). As convergence proceeds, the degree of correlation between X_0 and X_1 and increases. This has the effect of rapidly raising the condition number of $\mathbb{R}$. As a result, the determinant of $\mathbb{R}$ will fall, and its eigenvalues and singular values will become increasingly disparate.

[0043] Case 2: Converged, no double-talk

[0044] In this state $H_{0,0} \approx H_{1,0}$., so $X_0 \approx X_1$. This will make $\mathbb{R}$ very nearly rank deficient, and its condition number very large ($\kappa \sim 10^6$). Since $\mathbb{R}$ is close to being singular, its determinant will become very small. Similarly, we would expect to find only one significant eigenvalue or singular value.

8

[0045] <u>Case 3: Double-talk</u>

[0046] When double-talk is occurring, X_0 contains components from both S_0 and S_1, while X_1 is derived solely from S_0. In this case, X_1 and X_0 and are highly uncorrelated. R will have a low condition number, and this will be sustained for the duration of the double-talk. The higher the double-talk level, the lower the condition number becomes. This will raise the determinant of R, and we will find two significant eigenvalues and singular values.

[0047] Once the matrix R is generated, a variety of operations are available to determine what state the echo canceller is in. The condition number, determinant, eigenvalues and singular values of can all be used to test for double-talk or path changes. The determinant is used in the preferred embodiment because it is the simplest matrix operation to perform.

[0048] To illustrate the effectiveness of this algorithm at detecting double-talk and path changes, simulations were carried out and the results are shown in Figure 2. The plots indicate the value of det [R] under normal convergence, a path change, and double-talk. The scaling of the y-axis on the plots clearly demonstrates the variations observed in det [R] under the three different states. The simulations were carried out using ITU CSS synthetic speech signals from the G.168 Digital Echo Canceller standard. ITU-T Recommendation G.168, Digital Echo Cancellers. The signals were 48000 samples long, and a 60 ms echo path was used (which was changed to 15 ms during the path change simulation).

[0049] Under normal convergence (Fig. 2a), det [R] rapidly decays to near-zero values. When a path change occurs (Fig. 2b), det [R] spikes to a large value and then decays (to emphasize this trend, convergence was slowed by a factor of 10 following the path change). With double-talk (bottom plot), even larger, but sustained, spikes are present in det [R]. The differences in these three plots make it very easy to tell what state the echo canceller is in simply by checking the level of det [R]. The highest levels indicate double-talk, medium levels (along with decay) occurs with path changes, and very low levels are characteristic of full convergence. Based on these results, thresholds can be set as follows:

9

- Normal (converged) operation.

- Path change detected.

- Double-talk detected.

[0050] Once the state of the echo canceller is determined, the LMS filter operation can be adjusted accordingly.

[0051] A well-known relation in signal processing is that the convolution of two signals in time is equivalent to the multiplication of their frequency spectra. This property makes it possible to propose a variation on the preceding algorithm in which frequency-domain versions of the signals are used. X has been defined in the time-domain using convolutions:

$$[0052] \quad X[n] = H[n] \otimes S[n]$$

[0053] By taking the Fourier Transform of all terms involved, it is possible to rewrite the above equation in the frequency-domain as

$$[0054] \quad X(f_k) = H(f_k) \cdot S(f_k)$$

[0055] for all frequencies in the range $0 \leq f_k, \leq F_S/2$ where F_S is the sampling frequency of the signals. The generation and analysis of the correlation-based matrix $\mathbb{R}$ is carried out as before, only now $\mathbb{R}$ is created using the frequency-domain version of X.

[0056] The advantage to this approach is that the algorithm no longer needs to have an accurate echo path estimate for $\mathbb{R}$ to have a high condition number during non-double-talk states. The double-talk detector becomes completely insensitive to path changes. Depending on the application, this may or may not be a desirable property. Low-level double-talk detection abilities improve, but a separate path change detection scheme must now be used.

[0057] Implementation of a frequency-domain version of this process can be accomplished through the use of Fast Fourier Transforms (FFTs) or subbanding techniques.

[0058] As will be understood by persons skilled in the art the inventive process can be implemented in a digital signal processor or other suitable digital signal processing device.

5   [0059] Glossary

Adaptive Filter: A filter whose coefficients can be adjusted during operation. Adaptive filters are used to estimate unknown parameters, for example an unknown echo path.

Autocorrelation: A statistical quantity which roughly measures the similarity of a signal to time shifted versions of itself.

10   **Condition Number:** A measure of how close a matrix is to being singular. The condition number for an arbitrary matrix A is given by $\kappa(A) = \|A\|\|A^{-1}\|$.

Convergence: The condition achieved when the LMS filter has accurately modelled the echo path and is no longer undergoing significant changes. At convergence, the LMS filter is cancelling the maximum amount of echo.

15   **Cross-Correlation:** A statistical quantity which roughly measures the similarity of two separate signals.

Divergence: The process by which the LMS filter coefficients move away from the actual echo path to erroneous and unpredictable solutions. During divergence, the amount of echo being cancelled becomes less and less.

20   **Double-Talk:** The condition which occurs during simultaneous transmission of signals from both sides of the echo canceller.

Echo Path: A mathematical description of the process which imparts an echo to a signal.

ERL: Echo Return Loss. The loss a signal experiences as it travels along the echo path from R_{OUT} to S_{IN}.

25   **ERLE:** Echo Return Loss Enhancement. A common method of measuring the performance of an echo canceller. This measurement represents the amount that an echo signal has been reduced from S_{IN} to S_{OUT}.

LMS Algorithm: Least Mean Squares algorithm. Common adaptive filtering technique.

N-LMS Algorithm: Normalized Least Mean Squares algorithm. A variation on standard LMS in which the tap- weight update term is scaled by the inverse of the input signal power.

5   **Rank:** The number of non-zero eigenvalues or singular values a matrix has. Full-rank matrices have a non-zero determinant, and are thus non-singular and invertible.

RLS Algorithm: Recursive Least Squares algorithm. Common adaptive filtering technique.

It will be appreciated by one skilled in the art that many variations of the invention are
10   possible without departing from the scope of the appended claims.